D'CUMENT RESUME

ED 233 029 TM 830 367

AUTHOR Johnston, J. oward
TITLE Ethnography and Education: An Outsider's View of Ethnography and Educational Evaluation.
PUB DATE [83]
NOTE 9p.; For related document, see TM 830 366.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Credibility; Elementary Secondary Education; *Ethnography; Evaluation Methods; *Program Evaluation; *Reliability; Research Methodology; *Validity
IDENTIFIERS Observer Effect; Quantitative Research

ABSTRACT

The relevance of ethnographic research to evaluation of educational programs is discussed. The author focuses on evaluation questions that deal with finding out what is happening in a program and if what is happening is desirable. Ethnography suits the purposes of evaluators who answer such questions because they want to know how something works under the worst of conditions. Issues related to the validity and reliability of measurements used in evaluation illustrate this point. The internal validity of a study can be threatened by history, maturation, instrumentation, selection, mortality, and reactive arrangements--all part of the life of the school. To conduct an evaluation by controlling these variables is to study a version of the school setting that can never be replicated. Educational evaluations must be capable of focusing on the program as it really functions. Conventional measures of reliability in evaluations call for inter-subjective agreement. Ethnography, however, focuses on another form of reliability called "confirmability" where there is agreement from a number of information sources that a certain phenomenon has occurred. Finally, the author maintains that the success of an evaluation plan rests on its credibility. Naturalistic inquiry, as structured under ethnography, sustains credibility because of its logic, responsiveness, and involvement of participants. (LC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED233029

Ethnography and Education: An Outsider's View of Ethnography and Educational Evaluation

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

J. Howard Johnston, Ph.D.
University of Cincinnati

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. H. Johnston
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Professor Shultz's paper raises convincing theoretical and conceptual points that are especially relevant to the preparation of teachers and other education professionals. At the same time, though, he raises other issues that are germane to educational evaluation in general and the evaluation of educational programs in particular. In this brief paper, I will turn my attention to those issues.

Most evaluation is designed to find an answer to one of two kinds of questions. The first of these is: "What's happening here? Is it what we want?" The second is: "Is _____ happening here?" Clearly, the latter question is the one commonly answered through most existing quantitative approaches to evaluation. It assumes that a specific portion of an educational program can be isolated and studied independent of the other program elements around it. The difficulty is, however, that if the answer to the latter question is "no" we still don't know much about what is happening in the program we have studied.

TM 830 367

For that reason, I will focus on the former set of questions--those which deal with finding out what is happening in our program and if it is what we want to happen.

Evaluation has its own needs and issues quite apart from those normally discussed in research. Most of these issues focus on measurement and the quality of the measurement involved in the evaluation. Traditional concepts such as validity and reliability are especially worthy of note.

The validity question issue exists in two parts: internal validity and external validity. Internal validity deals with whether we really observed and measured what we thought we were observing and measuring. External validity deals with the extent to which what we find in our study is applicable to other settings and other times. Campbell and Stanley (1963) clarified internal validity in terms of the phenomena which threatens it: history, maturation, instrumentation, selection, mortality and reactive arrangements. Normally in experimental settings, we are concerned with the effects of history, the uncontrolled events which occur in the life of a school and may affect the study, and maturation, changes in performance that are brought about simply by the growth and development of individual students. Either of these events can produce outcomes not related to the treatment intervention; in our case, the educational program. In addition, observer effects are at issue because the observer (or instru-

ment) can actually interfere with the treatment being given so the results stem not from the treatment but from the instrument or the observer. Related issues include selection--where the uniqueness of an individual group's features produce spurious results or statistical regression produces gains which cannot be attributed to the treatment but to the extreme position in which the group began. Internal validity concerns itself also with experimental mortality--the effect on the experiment of having individuals enter or leave the experimental condition. Finally, reactive arrangements are those conditions accompanying an experimental condition but unrelated to it, and, thus confound the effects of the experimental treatment. Perhaps the most famous of these is the Hawthorne Effect.

Ironically, history, maturation, selection, regression, mortality, and reactive arrangements are all part and parcel of the life of the school. To conduct an evaluation by controlling these variables--by pretending that they do not exist in the life of schools--is not to study the school program at all, but to study some carefully controlled version of it which can never be replicated in a school setting. To illustrate, it is easy for us to imagine the surprise of a surgeon who says, "I don't understand it. The procedure worked fine in the clinic. Why did the patient die when we performed the surgery here in the forest?" Educational programs function in the forest, with all

of its dirt and lack of sterility, and educational program evaluations must be capable of focusing on both the procedure and the way it exists in the forest.

It must be granted that observer effects are more problemmatic than those noted above because of an evaluator's presence. This is especially true if we use Eisner's (1978) connoisseurship model where the evaluator is the instrument of evaluation. However, if we take Shultz's advice--and I argue that we should and make the participants (i.e., the teachers) the observers and the evaluators--, we are not really introducing an observer effect. That teachers are supposed to affect learning. Presumably we don't want to factor that out of the evaluation design; we want the teachers to continue helping children to learn. But rather than try to factor the the effects of the teacher out of our evaluation--a process we attempt all too often--we should instead factor out the intrusive effect of the evaluator from the program evaluation. The easiest way to do that is to make the teacher the evaluator as well.

We can argue that external validity is irrelevant in the case of program evaluation. External validity assumes that a program, and therefore its evaluation, can be translated to another setting and replicated. It simply cannot be. More important, it probably should not be reproducible in another setting. If one views the school as an intellectual/socio-

logical/psychological environment, it is quite impossible to imagine a case in which all of the variables associated with each of those dimensions exactly reproduce variables found in another setting. Evaluations are idiosyncratic; they must be related to the specific programs they assess, and, insofar as these programs are not interchangeable with others, neither are their evaluation designs should be.

Issues of reliability are important in educational evaluation plans. Conventional measures of reliability call for agreement among two or more competent observers or what Guba called "inter-subjective agreement." Another form of reliability, however, may be demonstrated by "confirmability" the case where there is agreement from a number of information sources that a certain phenomenon has occurred. In fact, if the kids, teachers, and parents don't understand the problems in the math book, the problems aren't understood, no matter how many math educators rate the materials as adequate to explain the relevant concepts. Conventional measures associated with program evaluation focus on the former type; ethnography focuses on the latter. Frankly, my preference is for the confirmability of data rather than simple inter-subjective agreements.

I have spent a lot of time talking about reliability and validity because at a meeting like this that is what one is expected to do. However, every educational evaluation plan on

which I have worked has stood or fallen on two issues: its credibility with its audience and its sensitivity to its social-political milieu.

Credibility can be translated into two specific questions. Does the audience believe the results and do they believe in the a process that produced them? The first question is one which suggests that the results should produce some intuitively positive reaction in the audience. They should fit with what our sensibilities tell us could be true. The second deals with the credibility of the researcher. Naturalistic inquiry has the distinct advantage of being tremendously logical. Researchers look at what is going on and study recurrent pieces of the event stream quite intensely. They look for patterns and constructs, and then, most important, they discuss what they think they have found with the participants in the process. Because it is so logical it does adhere to the collective knowledge base and belief systems of individuals in the situation. If we follow Shultz's advice, the participants are the process and are, therefore, are more likely to believe the results of the inquiry they have conducted.

The second question of credibility deals with the social political responsiveness of the evaluation. Evaluations are used to make decisions--decisions about resources, careers, learning opportunities, and programs. But the decisions themselves are

political and legislative. They are made by Boards of Education, by State Legislators, and in contract negotiations. Thus, evaluation is part of a political and social process. To pretend that it isn't by looking for clinical sterility is to isolate it even further from the reality of the schools and, thus, from its credible acceptance by the professionals in the schools.

Evaluative paranoia is induced when the participants know just how isolated the evaluation is from the political decision making process in which it must operate. Involving participants helps to reduce the paranoia by making clear that they are not being "studied;" in fact, they are studying themselves and the context in which they work. It makes sense, then, that the credibility of naturalistic inquiry, as structured under the general principles of Ethnography, sustain much more credible attention from the audience than do the more obtuse and abstract principals of much of our quantitative methodology.

In closing, it is worth pointing out Guba's remarks (1978) that the major purpose of evaluation is to respond to audience requirements for information. It is more important to improve audience understanding than to provide a "reductionist" summary. Indeed, the major criterion for validity in an evaluation is less "the truth value" of its judgments than the degree to which audience understanding is increased.

Ray Carpenter (in Guba, 1978) said that Ethnography invites interference... interference from the setting and all other "uncontrolled" variables. Ethnography, then, is especially well suited to our purposes as evaluators, for we want to know not how something works under the best of conditions but how it works under the worst of conditions. Because it is a way of thinking--a multidimensional way--it is important for teachers and other school professionals to learn the Ethnographic way of thinking in order to understand their world. I think Shultz is right. We should be teaching teachers and all school professionals this way of thinking about what they do.

References

Guba, E.G. Toward a Methodology of Naturalistic Inquiry in Educational Evaluation , Los Angeles: Center for Study of Evaluation, U.C.L.A., 1978.

Eisner, E. The Design and Evaluation of Educational Programs, New York: MacMillan, 1979.